

学校编码: 10384

分类号 \_\_\_\_\_ 密级 \_\_\_\_\_

学号: 18220051301722

UDC \_\_\_\_\_

厦 门 大 学

硕 士 学 位 论 文

声纹识别技术在司法鉴定中的应用研究

Voiceprint Recognition Technology and its Application in  
Forensic Identification

谢春荣

指导教师姓名: 张建寰 副教授

专 业 名 称: 精密仪器及机械

论文提交日期: 2008 年 5 月

论文答辩时间: 2008 年 月

学位授予日期: 2008 年 月

答辩委员会主席: \_\_\_\_\_

评 阅 人: \_\_\_\_\_

2008 年 5 月

厦门大学博硕士论文摘要库

## 厦门大学学位论文原创性声明

兹呈交的学位论文，是本人在导师指导下独立完成的研究成果。本人在论文写作中参考的其他个人或集体的研究成果，均在文中以明确方式标明。本人依法享有和承担由此论文产生的权利和责任。

声明人（签名）：

年 月 日

厦门大学博硕士论文摘要库

## 厦门大学学位论文著作权使用声明

本人完全了解厦门大学有关保留、使用学位论文的规定。厦门大学有权保留并向国家主管部门或其指定机构送交论文的纸质版和电子版，有权将学位论文用于非赢利目的的少量复制并允许论文进入学校图书馆被查阅，有权将学位论文的内容编入有关数据库进行检索，有权将学位论文的标题和摘要汇编出版。保密的学位论文在解密后适用本规定。

本学位论文属于

1. 保密（    ），在    年解密后适用本授权书。
2. 不保密（    ）

（请在以上相应括号内打“√”）

作者签名：                      日期：        年        月        日

导师签名：                      日期：        年        月        日

厦门大学博硕士论文摘要库

## 摘要

声纹识别技术属于生物认证技术的一种,声纹识别正是根据人的语音波形中反映说话人生理和行为特征的语音参数,自动识别说话人身份的技术。本文主要研究声纹识别技术在司法鉴定领域内的应用,包括对我国司法声纹鉴定的现状以及存在问题深入的探究和根据司法鉴定的要求进行声纹识别系统的开发。

目前,国内在司法声纹鉴定领域内存在很多问题,很多语音案件都无法被识别,其一个原因是我国现在只依靠大型的司法鉴定机构购买国外的价格昂贵的先进仪器来进行分析、鉴定。其二是我国在司法声纹鉴定方面没有形成行业标准。其三是汉语语言的独特性,即其发音复杂,方言多样性,以及音调的影响等,给我国声纹识别技术带来更大的困难和挑战。如果仅仅利用国外的声纹识别技术往往无法满足汉语语音的识别要求,本文致力于开发出低成本又能满足鉴定要求的声纹识别系统。

以满足司法声纹鉴定的技术需求和适应地方中小型鉴定机构资金能力为目的,本文在具体介绍了声纹识别技术的理论基础和分析了国内司法声纹鉴定的现状基础上做了下面几方面的工作:

- (1) 司法声纹鉴定研究。研究我国司法声纹鉴定存在的问题,论文根据声纹识别技术的要求,在鉴定程序和标准化建设方面进行深入的研究。
- (2) 预处理分析。由于司法案件中语音来源的复杂性,对原始语音信号的恢复,失真补偿,结合小波理论去除噪音,进行基音端点检测,滤除干扰等是本文的研究重点。
- (3) 特征参数的选取和提取。对比研究现有各种主流的语音信号特征提取算法,提取具有静、动态,更能反映说话人生理及行为的特征参数。如美尔倒谱系数(MFCC)及其反映动态特征的 $\Delta$ MFCC等特征参数。
- (4) 模式匹配算法及其系统的设计研究。重点分析了基于高斯混合模型的声纹识别确认系统。在 MATLAB6.5 平台上,结合当前最为流行的 Praat 软件,完成了特征提取,语音的 GMM 训练和识别,验证了声纹识别的理论以及本文所采用的分析方法,最后介绍本识别方法在案件中的应用。

**关键词:** 声纹识别; 司法鉴定; 高斯混合模型

厦门大学博硕士论文摘要库



## ABSTRACT

Voiceprint recognition (VR) technology, one of the biometric certification technologies, makes use of the speech coefficients in the waveform which represent the speaker's physiological and physical features to identify the speaker automatically. The paper mainly studies the application of voiceprint recognition in the forensic identification, including the status and problems of the forensic identification in China and the development for the VR system according to the demand of the forensic practice.

At present, there are three main reasons why more and more voice cases couldn't be identified. First, the bodies of forensic identification only depend costly on foreign advanced equipments to analyze and recognize voiceprint. Second, forensic voice identification is lacking of industry standardization in China. Finally, the uniqueness of the Chinese, the complex of pronunciation, dialect diversity, as well as the effect of tone, bring more difficulties and challenges to VR technology. If only utilizing foreign VR technology is often unable to meet the requirements of Chinese speech recognition. The paper aims to develop a low-cost system which also could meet the requirements of the forensic identification.

To meet the demands of forensic practice and adapt to the situation of shortage of funds in medium-sized forensic institutions, Based on the theoretical foundation of VR technology and the status of forensic voiceprint identification, the paper does some researches as follows:

- (1) Study the existing problems and reasons why more and more voice cases couldn't be identified. Conduct an in-deep research to procedures and standardization of forensic voice identification according to the demands of VR technology.
- (2) As the complexity of the original speech signal in the judicial case, the pretreatment analysis is very important. The key studies included in this paper, are the distortion compensation, de-noising based-on the wavelet theory, pitch endpoint detection and filtering interference

- (3) Extract static and dynamic features which better reflect what speaker physiological and behavioral parameters by studying and comparing the existing mainstream voice signal features, such as Mel frequency cepstrum coefficient (MFCC) and its difference ( $\Delta$ MFCC).
- (4) Study the pattern matching algorithm and system design. Mainly study the voiceprint verification system based on the Gaussian mixture model (GMM). The paper completed the feature extraction, GMM training and recognition by MATLAB6.5 and Praat software which also verified the theory of VR and the analysis methods. Besides, it also did some further analysis on the actual case of voiceprint recognition in forensic practice.

**Keywords:** Voiceprint Recognition; Forensic Identification; GMM

## 目录

<b>第一章 绪论</b>	<b>1</b>
1.1 课题研究背景	1
1.1.1 声纹识别技术的研究	2
1.1.2 声纹鉴定标准的建设	2
1.2 声纹识别的发展历史和研究现状	2
1.2.1 声纹识别技术的发展历史	2
1.2.2 国内外研究概况、水平及发展趋势	3
1.3 本论文主要工作和内容安排	4
<b>第二章 声纹识别技术概述</b>	<b>7</b>
2.1 声纹识别的理论依据	7
2.2 声纹识别的基本概念和分类	9
2.3 声纹识别的原理及关键技术	9
2.4 声纹识别技术的应用	11
2.5 本章小结	12
<b>第三章 司法声纹鉴定及其标准化建设</b>	<b>13</b>
3.1 我国司法鉴定的状况以及存在的问题	14
3.2 司法鉴定的标准化建设的迫切性	15
3.3 结论	17
<b>第四章 声纹的预处理及其特征提取</b>	<b>19</b>
4.1 语音信号的预处理	19
4.1.1 归一化处理	19
4.1.2 语音预加重	20
4.1.3 通道失真补偿	20
4.1.4 小波去噪处理	20
4.1.5 加窗	23
4.1.6 端点检测	24
4.2 声纹特征参数提取	27
4.2.1 基于倒谱法基频提取方法	27
4.2.2 基于线性预测的共振峰求取方法	29
4.2.3 LPCC 特征参数的提取	31
4.2.4 MFCC 特征参数提取	33
4.3 本章小结	37
<b>第五章 声纹识别中的模式匹配</b>	<b>39</b>
5.1 与文本相关的声纹识别的模式匹配	39
5.1.1 DTW 算法	39

5.1.2 DTW 基本原理 .....	40
<b>5.2 与文本无关的声纹识别的模式匹配 .....</b>	<b>41</b>
5.2.1 GMM 模型的基本概念 .....	41
5.2.2 GMM 模型的参数估计—EM 算法 .....	42
5.2.3 K 均值算法 .....	44
5.2.4 GMM 模型的识别 .....	45
5.2.5 GMM 模型的实际应用问题 .....	50
<b>5.3 基于径向基函数网络的声纹识别系统 .....</b>	<b>51</b>
<b>5.4 本章小结 .....</b>	<b>52</b>
<b>第六章 声纹确认系统的设计与应用 .....</b>	<b>53</b>
<b>6.1 系统开发环境介绍 .....</b>	<b>53</b>
6.1.1 硬件环境 .....	53
6.1.2 软件环境 .....	53
6.1.3 Praat 软件介绍 .....	54
<b>6.2 语音信号样本的采集 .....</b>	<b>55</b>
<b>6.3 语音信号的预处理 .....</b>	<b>56</b>
<b>6.4 实验的方案设计和结果分析 .....</b>	<b>58</b>
6.4.1 声纹的特征提取 .....	58
6.4.2 GMM 的训练实验 .....	58
6.4.3 识别结果和分析 .....	59
<b>6.5 声纹识别在司法鉴定中的应用程序探讨 .....</b>	<b>61</b>
<b>6.6 本章小结 .....</b>	<b>62</b>
<b>第七章 结论与展望 .....</b>	<b>63</b>
7.1 结论 .....	63
7.2 展望 .....	64
<b>[参考文献] .....</b>	<b>67</b>
<b>附 录 .....</b>	<b>70</b>
<b>致 谢 .....</b>	<b>76</b>
<b>硕士期间发表的论文 .....</b>	<b>77</b>

## CONTENTS

<b>CHAPTER 1 INTRODUCTION</b>	<b>1</b>
<b>1.1 Research background</b>	<b>1</b>
1.1.1 Research of VR technology	2
1.1.2 Building standard of voiceprint identification	2
<b>1.2 Research status and development direction of VR</b>	<b>2</b>
1.2.1 The development history of VR	2
1.2.2 Research status, level and direction	3
<b>1.3 Content and significance of research</b>	<b>4</b>
<b>CHAPTER 2 SUMMARY OF VOICEPRINT RECOGNITION</b>	<b>7</b>
<b>2.1 Theoretical basis of VR</b>	<b>7</b>
<b>2.2 Concept and classification of VR</b>	<b>9</b>
<b>2.3 Theory and key technology of VR</b>	<b>9</b>
<b>2.4 Application of VR</b>	<b>11</b>
<b>2.5 Conclusion</b>	<b>12</b>
<b>CHAPTER 3 FORENSIC'S VR AND STANDARDIZATION BUILDING</b>	<b>13</b>
<b>3.1. Status and problems of forensic's VR in China</b>	<b>14</b>
<b>3.2 Urgency of standardization building in forensic's VR</b>	<b>15</b>
<b>3.3 Conclusion</b>	<b>17</b>
<b>CHAPTER 4 PRETREATMENT AND FEATURE EXTRACTION OF VOICEPRINT</b>	<b>19</b>
<b>4.1 Pretreatment of speech signal</b>	<b>19</b>
4.1.1 Unitary processing	19
4.1.2 Pre-aggravation of speech	20
4.1.3 Channel's distortion and compensation	20
4.1.4 De-noise processing using wavelet transforms	20
4.1.5 Windowed	23
4.1.6 Endpoint detection	24
<b>4.2 Parameters extraction of voiceprint feature</b>	<b>27</b>
4.2.1 Pitch frequency extraction based on CEP	27
4.2.2 Formant extraction based on Linear Prediction	29
4.2.3 LPCC extraction	31
4.2.4 MFCC extraction	33
<b>4.3 Conclusion</b>	<b>37</b>
<b>CHAPTER 5 PATTERN MATCHING OF VR</b>	<b>39</b>
<b>5.1 Text-dependment speaker recognition of pattern matching</b>	<b>39</b>
5.1.1 DTW algorithm	39
5.1.2 Fundamental of DTW	40
<b>5.2 Text-independment speaker recognition of pattern matching</b>	<b>41</b>

5.2.1 Basic concept of GMM .....	41
5.2.2 Parameter estimate of GMM—EM algorithm .....	42
5.2.3 K-mean algorithm .....	44
5.2.4 GMM recognition .....	45
5.2.5 Problems of practical application of GMM .....	50
<b>5.3 VR system based on RBFN .....</b>	<b>51</b>
<b>5.4 Conclusion. ....</b>	<b>52</b>
<b>CHAPTER 6 DESIGN AND APPLICATION OF VOICEPRINT</b>	
<b>VERIFICATION SYSTEM .....</b>	<b>53</b>
<b>6.1 Introduction environment of system development .....</b>	<b>53</b>
6.1.1 Hardware environment.....	53
6.1.2 Software environment .....	53
6.1.3 Praat introduction.....	54
<b>6.2 Voice acquisition. ....</b>	<b>55</b>
<b>6.3 Voice pretreatment. ....</b>	<b>56</b>
<b>6.4 Project design and result analysis of experiment. ....</b>	<b>58</b>
6.4.1 Feature extraction of voiceprint .....	58
6.4.2 Training experiment of GMM.....	58
6.4.3 Result and analysis.....	59
<b>6.5 Application in forensic identification practice. ....</b>	<b>61</b>
<b>6.6 Conclusion. ....</b>	<b>62</b>
<b>CHAPTER 7 CONCLUSION AND OUTLOOK .....</b>	<b>63</b>
<b>7.1 Conclusion. ....</b>	<b>63</b>
<b>7.2 Outlook .....</b>	<b>64</b>
<b>[REFERENCES] .....</b>	<b>67</b>
<b>APPENDIX .....</b>	<b>70</b>
<b>THANKS .....</b>	<b>76</b>
<b>PUBLICATIONS .....</b>	<b>77</b>

## 第一章 绪论

基于生物特征的身份识别技术是当前国际上的重点研究内容，声纹识别是生物特征识别的一种，是通过语音识别来确认说话人的身份。在系统安全认证、司法鉴定、金融服务以及电子侦听等领域有着广泛的应用价值，而且比起其他生物特征识别技术如指纹识别、脸部识别等，更加经济简便，特别是对于远距离身份确认更为有效。

### 1.1 课题研究背景

由于我国司法机关内设司法鉴定机构规模大，鉴定专业技术人员较多，鉴定仪器设备基本来自国外，是当前传统的司法鉴定设备的主力军<sup>[1]</sup>，如司法部司法鉴定科学技术研究所，北京华夏物证中心，法大法庭科学技术鉴定研究所，华东政法学院司法鉴定中心都可以进行声纹鉴定，但他们一般采取国外昂贵的大型声纹鉴定分析仪器进行鉴定，如 KAY DSP 5500 语图仪(Sona-Graph)，其最大弊端不只因为它价格高，而且不适合语音的弹性分析，导致国内大量语音案件积压而没办法处理。

为了充分利用司法鉴定资源和方便人民群众的诉讼活动，在地方司法鉴定法规和部门规章中允许司法机关内设鉴定机构并开展对外司法鉴定服务活动。原来的司法鉴定全部由司法部门自己做鉴定，现在改革为由独立的鉴定机构，做为司法程序中第三方机构出现。本课题的目的是开发出适用于中小型司法鉴定机构的声纹识别系统，是为适应与福建正泰司法鉴定中心进行相关而开展的研究项目，福建正泰司法鉴定中心是厦门市首批获得福建省司法厅批准成立的司法鉴定机构，属于社会司法鉴定机构，其力量相对薄弱，对语音声纹鉴定技术的需求非常迫切，根据国内目前声纹识别技术现状及法律适用情况，拟从声纹识别技术和声纹鉴定标准两方面开展工作。

### 1.1.1 声纹识别技术的研究

生物特征识别技术早年运用于法庭科学的司法鉴定,多是对静态图像<sup>[2][5]</sup>(指纹图像、脸形图像)的事后采集和识别;上世纪九十年代末和本世纪初,特别是9.11恐怖事件之后,由于国际反恐斗争的需要,对静态、动态特征的事前事后采集和实时鉴别,已成为防范安全风险的主要技术手段。同时,由于汉语语音和其他语言存在明显的不同,仅依靠国外先进的设备很难满足国内的语音案件的处理。因此,建立一套适合我国的声纹识别系统是国内研究该领域的学者的重要任务。声纹识别技术研究在实验室环境中已经得到很大的发展,但由于实际环境中影响语音识别的因素很多,实际应用中还不够广泛,所以研究在实际应用中的推广是很有必要,也是很迫切的,基于司法鉴定领域中的声纹识别研究就显得很有必要,也是安全领域里面急需解决的课题。

### 1.1.2 声纹鉴定标准的建设

当前国内积压了很多只存有语音资料的司法案件还没办法判定,其一个重要原因是我国在司法声纹鉴定采用不规范的传统标准<sup>[1]</sup>。我国的公共安全行业是国内最早应用人体生物特征识别技术的行业,也是最早制定人体生物特征识别技术标准的行业。各类基于指纹、掌形、面像识别的技术和产品、系统已广泛应用于刑事侦察、法定证件制作、出入境人员管理、大型国内国际会议的出入口控制等公安业务和安全防范工作中。但由于声音非接触性、动态特征等特性,声纹司法鉴定标准的建设面临更大的挑战。

## 1.2 声纹识别的发展历史和研究现状

### 1.2.1 声纹识别技术的发展历史

对声纹识别的研究始于20世纪30年代。早期的工作主要集中在人耳听辨试验和探讨听音识别的可能性方面。随着研究手段和工具的改进,研究工作逐渐脱离了单纯的人耳听辨。Bell实验室的劳伦斯·科斯塔(L.G. Kesta)目视观察语谱图进行识别,提出了“声纹”的概念<sup>[3]</sup>。之后,电子技术和计算机技术的发展,使通



Degree papers are in the "[Xiamen University Electronic Theses and Dissertations Database](#)". Full texts are available in the following ways:

1. If your library is a CALIS member libraries, please log on <http://etd.calis.edu.cn/> and submit requests online, or consult the interlibrary loan department in your library.
2. For users of non-CALIS member libraries, please mail to [etd@xmu.edu.cn](mailto:etd@xmu.edu.cn) for delivery details.

厦门大学博硕士论文摘要库